Claims

What is claimed is:

1. A computer-implemented method for comparing two databases which each comprise Chinese text data items specifying addresses, the method comprising:

for each of the databases, converting any of the Chinese text data items which are not in a predefined common Chinese language format into that common format, any items in at least a first of the databases which are convertable into the common format in multiple ways being converted in all those ways; and

comparing the data items in the common format, to identify Chinese text data items in the first database corresponding to Chinese text data items in the second database.

2. The method of claim 1 wherein the common Chinese language format is Pin Yin characters.

3. The method of claim 1 wherein the first database is a database of an order management system and includes data items including a corresponding shipping and/or billing addresses.

4. The method of claim 1 wherein the Chinese text data items of the first database are in Simplified Mandarin characters.

5. The method of claim 1 wherein the second database is a list of proscribed parties issued by a governmental organization.

6. The method of claim 1 wherein data items of the second database which do

2   not include an address within at least one specified Chinese territory are not

3   converted into the common format.

1

1   7.   The method of claim 1 wherein upon the comparison of the Chinese text data

2   items in the common format a match between them is found, the corresponding item

3   of the first database is removed.

1

1   8.   A computer system for comparing two databases which each comprise

2   Chinese text data items specifying addresses, the computer system comprising:

3   a first conversion unit for converting the Chinese text data items of a first of

4   the databases into a predefined common Chinese language format, any items in a

5   first database which are convertable into the common format in multiple ways being

6   converted in all those ways to generate items in the common format;

7   a second conversion unit for converting the Chinese text data items of the

8   second of the databases into the common Chinese language format, and

9   a comparison unit for comparing the converted data items to identify the

10   Chinese text data items in the first database which correspond to Chinese text data

11   items in the second database.

1

1   9.   The system of claim 8 wherein the common Chinese language format is Pin

2   Yin characters.

1

1   10.   The system of claim 8 wherein the first database is a database of an order

2   management system and includes Chinese text data items including a

3   corresponding shipping and/or billing address.

1

1   11.   The system of claim 8 wherein the Chinese text data items of the first

2   database are in Simplified Mandarin characters.

1

1 12.   The system of claim 8 wherein the second database is a list of proscribed

2 parties issued by a governmental organization.

1

1 13.   The system of claim 8 wherein the second conversion unit does not convert

2 data items of the second database which do not include an address within at least

3 one specified Chinese territory.

1

1 14.   The system of claim 8 wherein the comparison unit is arranged, upon

2 detecting a match between the Chinese text data items in the common format, to

3 operate output means to signal that the corresponding item of the first database

4 should be removed.

1

1 15.   The system of claim 8 wherein the comparison unit is arranged, upon

2 detecting a match between the Chinese text data items in the common format, to

3 remove the corresponding item from the first database.